JOURNAL OF APPLIED BEHAVIOR ANALYSIS 2005, **38,** 469–484 NUMBER 4 (WINTER 2005)

EFFECTS OF STIMULUS VARIATION ON THE REINFORCING CAPABILITY OF NONPREFERRED STIMULI

LEAH J. KOEHLER AND BRIAN A. IWATA

UNIVERSITY OF FLORIDA

EILEEN M. ROSCOE

NEW ENGLAND CENTER FOR CHILDREN

AND

NATALIE U. ROLIDER AND LAURA E. O'STEEN

UNIVERSITY OF FLORIDA

We examined the effects of stimulus (reinforcer) variation in several different contexts. In Study 1, we identified high-quality (HQ) and low-quality (LQ) stimuli based on results of a paired-stimulus assessment and examined their effects when available under concurrent-reinforcement schedules for 8 participants. No participants showed preference for the LQ stimuli when compared singly or in a varied arrangement to the HQ stimulus. In Study 2, we identified nonpreferred (NP) stimuli based on results of a single-stimulus assessment and examined their effects when available under single-reinforcement schedules for 3 participants. Results of Study 2 were mixed. One participant's data indicated that the varied presentation of NP stimuli produced a modest improvement in performance over that observed when the stimuli were presented singly. By contrast, a second participant's data showed no facilitative effect for the varied delivery of NP stimuli and that the inclusion of an HQ stimulus in the varied arrangement obscured the reinforcing effects of the HQ stimulus. The 3rd participant's data showed no effect for the varied delivery of NP stimuli but an apparent facilitative effect when an HQ stimulus was included in the varied arrangement, which was attributable solely to the presence of the HQ stimulus.

DESCRIPTORS: preference, reinforcer variation

The identification of reinforcers is an important prerequisite to the development of effective behavioral contingencies, and a great deal of research has shown that empirical methods for assessing preference can be very helpful in selecting reinforcers for individuals with severe disabilities (Ivancic, 2000). A number of preference-assessment procedures have been described in the literature; commonly used methods include those in which approach responses are observed to stimuli presented singly (Pace, Ivancic, Edwards, Iwata, & Page, 1985), in pairs (Fisher et al., 1992), or in multistimulus arrays (DeLeon & Iwata, 1996).

Empirical reinforcer identification does not necessarily guarantee positive treatment outcomes because a number of factors may determine whether a reinforcement effect is produced or maintained. One of these factors is repeated exposure to the same reinforcer, which may result in performance decrements that often are attributed to satiation or habituation. Egel (1980, 1981) conducted a series of studies showing that this problem may be alleviated by varying the stimuli that are delivered as reinforcers. Egel (1980) observed that the pre-

This study was based on a thesis submitted to the University of Florida Psychology Department in partial fulfillment of the requirements for the MS degree and was supported in part by a grant from the Florida Department of Children and Families. We thank Sarah Bloom and Nicole Hausman for their assistance in conducting the study.

Reprints may be obtained form Brian Iwata, Psychology Department, University of Florida, Gainesville, Florida 32611.

doi: 10.1901/jaba.2005.102-04

sentation of varied reinforcers produced higher response rates than did the presentation of any one of the reinforcers singly. He subsequently extended those findings to classroom performance (Egel, 1981).

An interesting implication of Egel's (1980, 1981) findings is that varied presentation of less preferred reinforcers might compete with a more preferred reinforcer. Bowman, Piazza, Fisher, Hagopian, and Kogan (1997) addressed this question by first establishing a hierarchy of preference for a number of stimuli. They subsequently compared, under a concurrent-schedule arrangement, the reinforcing effects of the highest ranked (high-quality or HQ) stimulus presented singly versus those ranked second, third, and fourth (slightly lower quality or SLQ) presented in a varied format. Results showed that 2 of the 7 participants preferred the single HQ reinforcer over the varied SLQ reinforcers, 3 participants preferred the SLQ reinforcers, and the remaining 2 participants showed switches in preference across sessions. Thus, the SLQ stimuli disrupted a strong preference for the HQ stimulus in 5 of the 7 participants (i.e., the 3 who preferred the SLQ stimuli and the 2 who showed switches in preference). In light of these somewhat mixed findings and the fact that the rankings of the HQ and SLQ stimuli were very similar (in some cases, one or more of the SLQ stimuli had preference rankings identical to that of the HQ stimulus), the extent to which variation produces competition between less and more preferred reinforcers remains unknown. The purpose of Study 1 was to replicate the Bowman et al. procedures using HQ and low-quality (LQ) stimuli whose initial preference rankings showed a greater degree of disparity.

A further implication of findings from research on reinforcer variation is that the inclusion of an HQ stimulus as one of the consequences in a varied array that contains nonpreferred stimuli might actually strengthen the reinforcing effectiveness of stimuli that otherwise would not function as reinforcers. We examined this possibility in Study 2 by evaluating the effects of LQ stimuli when used as reinforcers alone, under a varied arrangement, and under a varied arrangement that included one HQ stimulus.

STUDY 1

Method

Participants and Setting

Eight individuals with developmental disabilities participated. Josh was a 31-year-old man who had been diagnosed with mild mental retardation and cerebral palsy. Matt was a 24-year-old man who had been diagnosed with mild mental retardation. Julie was a 41-year-old woman who had been diagnosed with profound mental retardation and Down syndrome. Bill was a 24-year-old man who had been diagnosed with mild mental retardation. Sam was a 25-year-old man who had been diagnosed with moderate mental retardation, autism, and Down syndrome. Tom was a 37-year-old man who had been diagnosed with mild mental retardation and Klinefelter's syndrome. Peter was a 27-year-old man who had been diagnosed with moderate mental retardation. Jack was a 37-year-old man who had been diagnosed with mental retardation and seizure disorder. All participants were ambulatory and could follow some instructions and communicate vocally, except for Julie, who could communicate with limited gestures only.

Sessions were conducted at a vocational day program that all participants attended. Reinforcer assessment sessions were 10 min in duration and were conducted three to four times daily, 4 to 5 days per week.

Response Measurement and Reliability

During the preference assessments, observers recorded, on a preprinted data sheet, the occurrence of approach responses to one of two edible items. An approach response was defined as the participant reaching for or

picking up one edible item within 5 s of the verbal prompt "pick one." Preference for each stimulus was calculated by dividing the number of trials on which a stimulus was approached by the number of trials on which it was presented. An independent observer collected data on 57% (range, 50% to 85%) of sessions. Reliability was calculated on a trial-by-trial basis by dividing the number of agreements by the total number of trials and multiplying by 100%. Mean reliability for the preference assessments was 99% (range, 97% to 100%).

During reinforcer assessment sessions, observers using laptop computers recorded data on the frequency of target behaviors and the delivery of reinforcers. An independent observer collected data on 34% (range, 29% to 41%) of all sessions. Session time was divided into continuous 10-s intervals, and observers' records were compared on an interval-by interval basis. Interobserver agreement was calculated by dividing the smaller number of responses in each interval by the larger number, averaging these fractions, and multiplying by 100%. Mean interobserver agreement was 98% (range, 94% to 100%) for all responses (HQ, LQ, SLQ, and control). Reliability was also calculated for the delivery of reinforcement and averaged 95% (range, 91% to 97%).

Preference Assessment

A paired-stimulus preference assessment (Fisher et al., 1992) was conducted for each participant. Sixteen edible items were selected based on results of informal interviews with staff members familiar with the participant. The participant was allowed to sample each edible item prior to the start of the assessment. On each trial, two items were presented, and the participant was allowed to select and consume one of the items (simultaneous approach to both items was blocked). Each item was paired twice with every other item during the assessment, and pairs were presented in a semi-random order (no pair was presented twice consecutively). An HQ stimulus was defined as one selected on at least 85% of the trials, an LQ stimulus was defined as one selected on 30% or fewer trials, and an SLQ stimulus was defined as one that was ranked second, third, or fourth in the preference assessment.

Experimental Conditions

Initial training. Each participant was taught to perform a target task prior to the reinforcer assessment. Target behaviors were selected for each participant based on three criteria: (a) The response was discrete and easily measurable, (b) it consisted of or approximated a vocational task, and (c) it could be performed by the participant. If a vocational task that the participant could learn quickly was not identified, a simple, arbitrary task (e.g., switch pressing) was selected. If the participant did not emit the response following a modeled prompt, brief training (gestural and physical prompting) was conducted until the participant did so independently. During the first stage of training, occurrences of the target behavior were followed by the delivery of an edible item on a fixed-ratio (FR) 1 schedule. The item used during training had been selected on 40% to 60% of the trials during the preference assessment and was different than those used during subsequent conditions. During Stage 2, a similar task (e.g., with materials of a different color) was available concurrently but produced no consequences. The reinforcer to be delivered was placed on a plate behind one of the tasks, and its position was changed following every few responses. Training continued until the participant consistently performed the task associated with reinforcement. During subsequent conditions (see below), three tasks (identical except for color) were concurrently available. Responses on either of two tasks (A or B) produced consequences (placed on plates behind the tasks); responses on the third task (C) produced no consequences and served as the control. The control task was included to identify situations in which participants responded randomly or failed to make discrimi-

nations. The response–stimulus pairings were randomized across sessions. Responses on the control task were rare, so these data are not presented in the results but are available from the authors.

Single HQ versus single LQ stimuli (HQ vs. LQ1, LQ2, LQ3). The purpose of this condition was to verify preference for the HQ stimulus over each of the LQ stimuli. This comparison was critical because if any of the LQ stimuli competed with the HQ stimulus, any varied condition containing that LQ stimulus would be unnecessary. In three sequential comparisons, the reinforcing efficacy of the HQ stimulus was compared with that of each LQ stimulus. The stimuli were available in a concurrent FR1 FR1 arrangement: Each response to the HQ task produced access to the HQ stimulus; each response to the LQ task produced access to the LQ stimulus, and each response to the control task produced no consequences. One plate was situated behind each task and contained the edible items associated with that task. The participant was prompted to engage in the target response associated with the HQ stimulus, the LQ stimulus, and the control prior to the start of each session.

Single HQ versus VLQ stimuli (HQ vs. VLQ). The purpose of this condition was to determine whether the LQ stimuli, when presented in a varied format, would compete with the single HQ stimulus. Procedures were similar to those in the previous condition, except that each response to the LQ task produced access to one of the three LQ stimuli. All of the varied items were visible to the participant throughout the session, and the order of stimulus delivery was fixed within a given session but randomized across sessions. The participant was prompted to engage in each target response three times prior to a session to establish contact with each available reinforcer.

Single HQ versus varied SLQ stimuli (HQ vs. VSLQ). The purpose of this condition was to determine whether the SLQ stimuli, when presented in a varied format, would compete with the single HQ stimulus. Procedures were similar to those in the previous condition, except that each response to the SLQ task produced access to one of the SLQ stimuli.

Experimental Design

These conditions were arranged in sequential order. Within each condition, a concurrent schedule was used to assess preference for one of the reinforcer options.

Results and Discussion

Figure 1 shows results of the preference assessments. The item selected as the HQ stimulus was chosen most frequently by each participant and on 100% of the trials by most participants. The only exception to this was Tom: His highest ranked item (Cuddlefish Peanut) was unavailable for purchase when we began his reinforcer assessment, so we used his second-ranked stimulus as the HQ stimulus. The three items ranked just below the HQ stimulus were selected as SLQ stimuli. Finally, the three lowest ranked items were selected as LQ stimuli.

Figure 2 shows results from the reinforcer assessments obtained for Josh (folding socks), Matt (copying words from note cards), Julie (operating a hole puncher), and Bill (copying words from note cards). When the HQ stimulus was compared to each of the three LQ stimuli individually, all 4 participants allocated more responding to the task associated with the HQ stimulus. In most cases, participants never made any responses to the task associated with the LQ stimulus (Matt was the exception). When the HQ stimulus was compared with the three LQ stimuli presented in a varied format, Josh showed frequent switches between the HQ and VLQ tasks. By contrast, Matt, Julie, and Bill continued to show preference for the HQ task; they never responded on the VLQ task. When VSLQ stimuli were substituted for VLQ stimuli, all participants allocated more responding to the VSLQ task. Josh and Matt showed

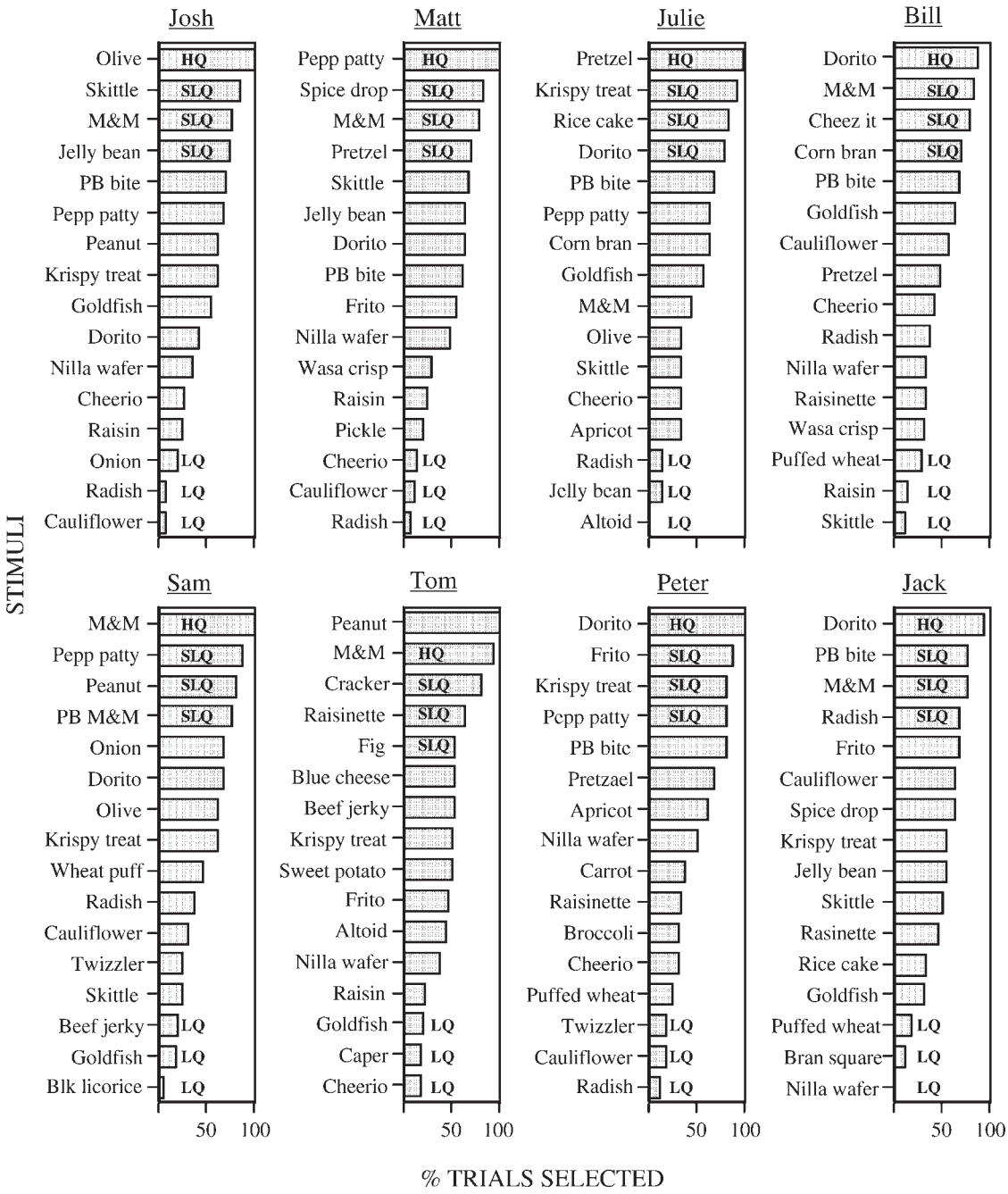

Figure 1. Results of paired-stimulus preference assessments. Items designated as high quality (HQ), slightly lower quality (SLQ), and low quality (LQ) were used in Study 1.

exclusive preference for the VSLQ task; Julie and Bill switched periodically between the HQ and VSLQ tasks but showed a general preference for the SLQ task.

Figure 3 shows results from the reinforcer assessments obtained for Sam (pressing a button), Tom (stuffing envelopes), Peter (folding socks), and Jack (punching holes in cards and

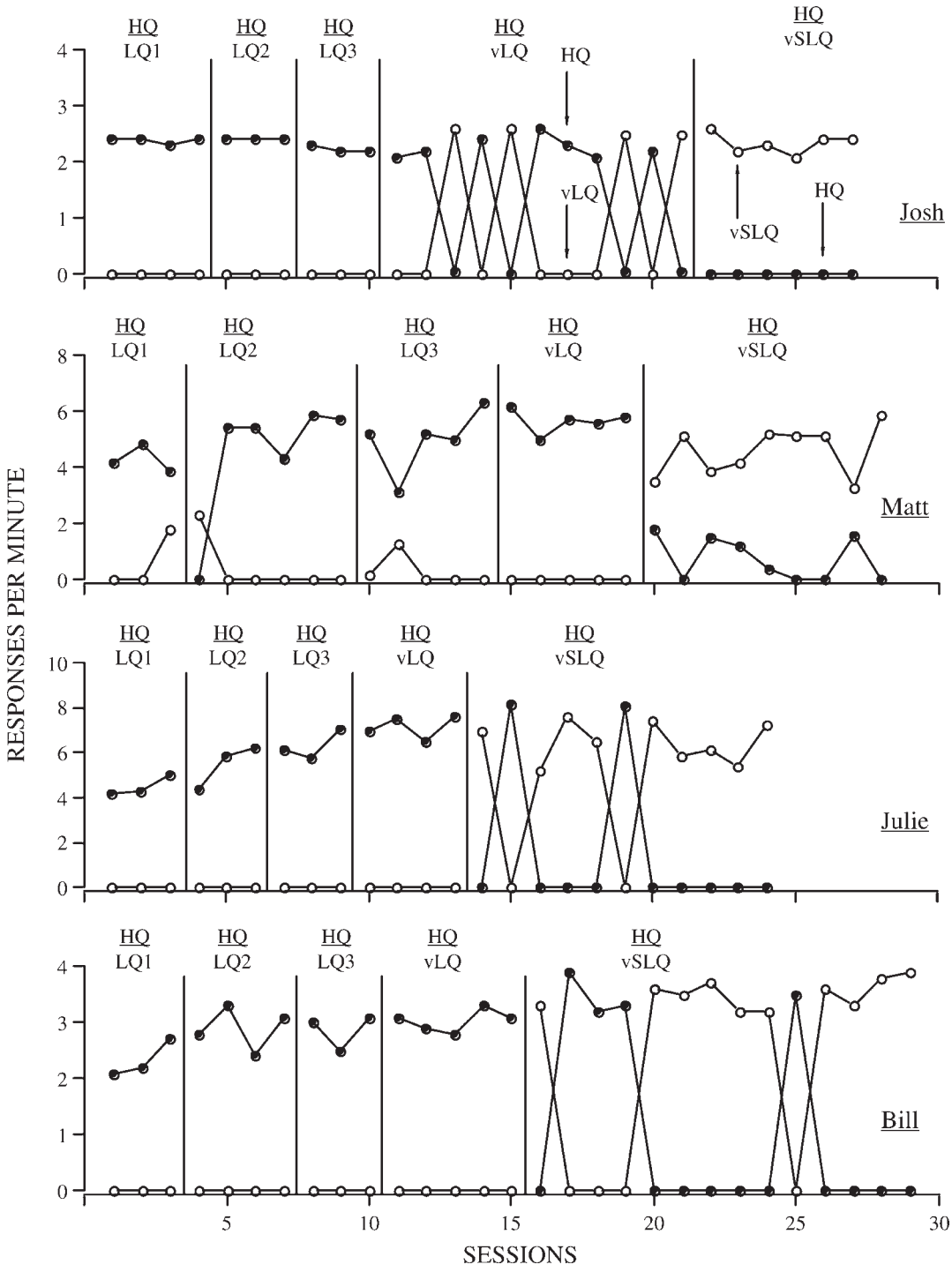

Figure 2. Results of reinforcer assessments for Josh, Matt, Julie, and Bill.

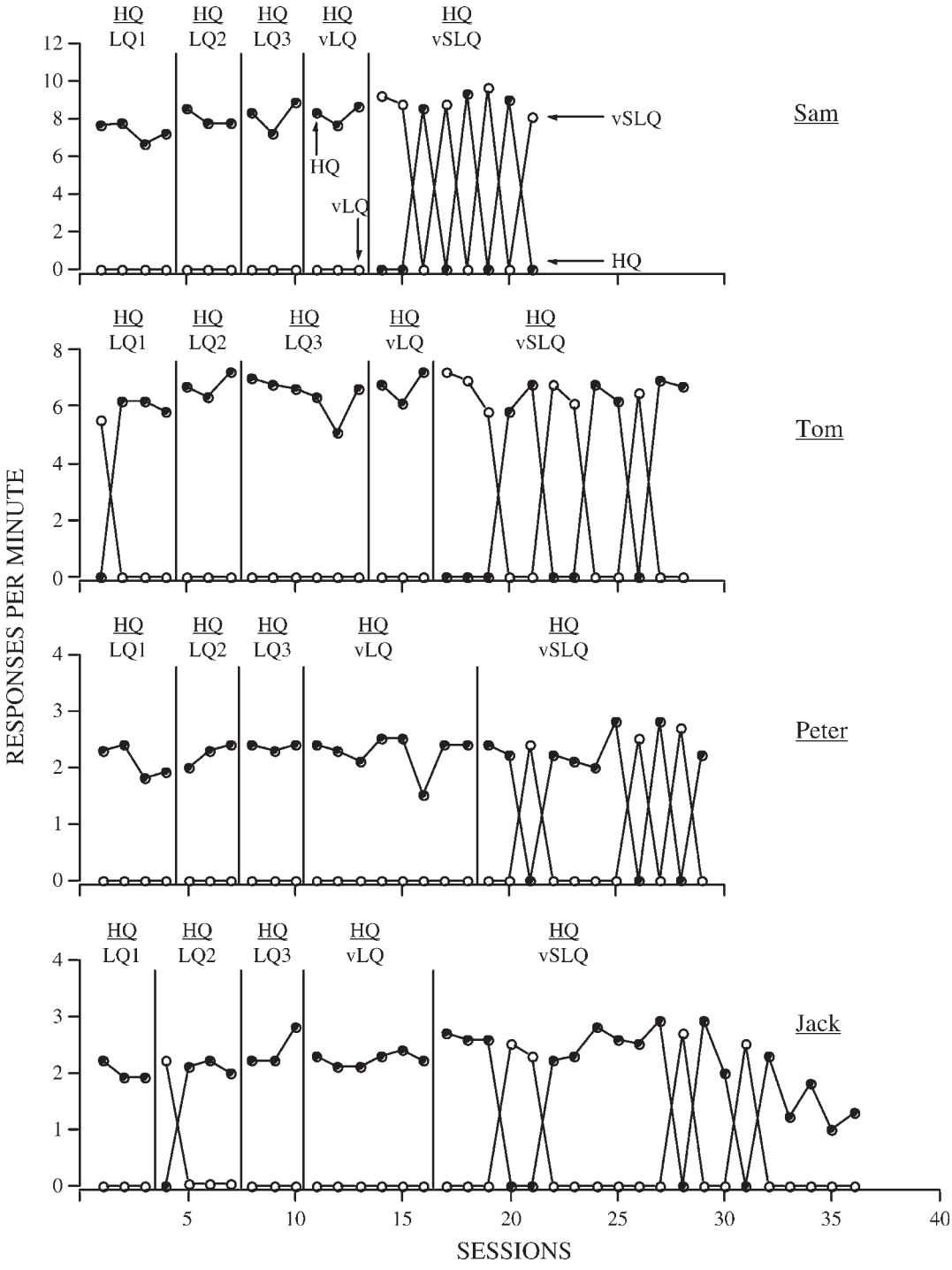

Figure 3. Results of reinforcer assessments for Sam, Tom, Peter, and Jack.

putting them in envelopes). All 4 participants showed preference for the HQ task over the LQ task when any one of the LQ stimuli was presented alone and when all three LQ stimuli were presented in a varied format. During the final condition, when the HQ stimulus was compared with VSLQ stimuli, all participants showed switching between the two response options, although Jack's performance tended to favor the HQ option. On 16 of 20 sessions in Jack's HQ versus VSLQ condition, he allocated more responding to the HQ task.

In summary, all participants showed preference for the HQ stimulus over each of the single LQ stimuli when both were concurrently available, which was a prerequisite for subsequent comparisons. When the reinforcing effects of the HQ stimulus were compared to those of VLQ stimuli, none of the participants showed preference for the VLQ consequences. Josh was the only participant whose performance showed some switching between the HQ and VLQ conditions. The consistency of these results indicates that stimulus variation does not appear to strengthen the reinforcing effects of low-preference stimuli when highly preferred stimuli are available concurrently.

By contrast, when the reinforcing effects of the HQ stimulus were compared to those of varied stimuli whose initial preference rankings were slightly lower than that of the HQ stimulus (VSLQ), half of the participants (Josh, Matt, Julie, and Bill) showed preference for the varied condition; the remaining participants (Sam, Tom, Peter, and Jack) switched between reinforcement options (i.e., they showed no clear preference). These results were similar to those reported by Bowman et al. (1997), who observed that 3 of 7 participants preferred the VSLQ condition. It is unclear, however, whether the shift in preference observed for some of the participants in our study, as well as in the Bowman et al. study, was a function of stimulus variation per se. Because neither study included a condition in which the HQ stimulus was compared to single SLQ stimuli (i.e., analogous to the first three conditions of our reinforcer assessment), the relative reinforcing effects of stimuli with similar preference rankings were unknown. Thus, it is possible that, under a concurrent arrangement, some participants would have shown preference for a single SLQ stimulus over the HQ stimulus or, at the very least, some switching across sessions.

Two additional features of the data warrant discussion. First, the two general patterns of responding (preference for HQ during the HQ vs. VLQ condition; switching during the HQ vs. VSLQ condition) were not observed for every participant (see Josh and Matt), and it is likely that these patterns were influenced by either the specific stimuli or the number of stimuli included in the initial preference assessments. In other words, although it seemed reasonable to define preference (HQ, SLQ, LQ) based on rankings, a given ranking may not be predictive of a reinforcement effect (Roscoe, Iwata, & Kahng, 1999). Depending on which and how many stimuli are included in a preference assessment, it is possible that few stimuli, all stimuli, or all stimuli above a given rank would function as reinforcers. Thus, the difference in results observed between the VLQ stimuli and VSLQ stimuli (when compared to the HQ stimulus) must be interpreted with caution.

Second, although the HQ versus VSLQ condition yielded switches in preference for 5 participants, most of the switching occurred across sessions; by contrast, most participants consistently chose either the HQ stimulus or the VSLQ stimuli (i.e., they did not switch) within a given session. Although positioning of the reinforcers changed across sessions, none of the participants was observed to show a position bias. The effects of satiation do not seem to account for this effect unless satiation to one reinforcer (but not another) occurred across sessions but not within a session. The finding that subjects often respond exclusively toward

one option under concurrent FR schedules (Herrnstein & Loveland, 1975) also seems unlikely as an explanation because it assumes that one of the schedules was more favorable, and it is not clear how one schedule would be more favorable on some sessions but not on others. It is, of course, possible that other variables, such as the type or amount of food consumed outside the session, may influence either satiation or the favorability of a given reinforcer on a given day (e.g., North & Iwata, 2005). In the present study, access to the stimuli used as reinforcers was controlled outside the sessions, and sessions were conducted at consistent times relative to meals, but type and amount of food were otherwise uncontrolled.

STUDY 2

Study 1 examined the effects of stimulus variation in the context of reinforcer competition under a concurrent-schedule (HQ vs. VLQ) arrangement. In applied situations, however, therapists and educators may be more concerned with whether stimulus variation enhances the effects of reinforcement when they attempt to establish or maintain a particular (single) performance, and the concurrent schedule used in Study 1 does not address this issue. Using a single-schedule arrangement, Egel (1980, 1981) showed that stimulus variation produced better results than those obtained with the same reinforcers delivered singly; however, the single-stimulus conditions initially showed positive results. Thus, all the stimuli functioned as reinforcers. As an attempt to extend Egel's research, in Study 2 we sought to determine whether stimulus variation might strengthen the reinforcing effects of stimuli that do not function initially as reinforcers. If so, therapists might actually be able to expand the range of reinforcers by varying the delivery of a highly preferred consequence with the delivery of consequences for which preference is negligible.

Method

Participants and Setting

Three participants with developmental disabilities participated. Ed was a 47-year-old man who had been diagnosed with moderate mental retardation and autism. Dan was a 27-year-old man who had been diagnosed with moderate mental retardation. Brian was a 44-year-old man who had been diagnosed with mild mental retardation. All participants were ambulatory, could follow instructions, and communicated vocally. The setting and session schedule were the same as in Study 1.

Response Measurement and Reliability

During the preference assessments, observers recorded the occurrence of approach responses as in Study 1. An approach response was defined as the participant reaching for or picking one edible item within 5 s of the verbal prompt "pick one" (paired-stimulus assessment) or picking and then consuming the edible item being placed in front of the participant (single-stimulus assessment). An independent observer collected data on 33% (range, 0% to 100%) of trials during the paired-stimulus preference assessment and on 100% of the trials during the single-stimulus preference assessments. Reliability was calculated in the same manner as in Study 1 and was 100%.

During reinforcer assessment sessions, observers using laptop computers collected data on the frequency of target behaviors and the delivery of reinforcers. An independent observer collected data during 37% (range, 30% to 44%) of all sessions. Reliability was calculated in the same manner as in Study 1. Mean interobserver agreement was 98% (range, 85% to 100%) for the target response and 99% (range, 85% to 100%) for the delivery of reinforcers.

Preference Assessment

A paired-stimulus assessment was conducted to identify an HQ stimulus using the same criterion that was used in Study 1. A single-stimulus preference assessment (Pace et al.,

1985) was conducted to identify nonpreferred (NP) stimuli. The rationale for using this assessment was that stimuli not selected in a paired-stimulus format may nevertheless function adequately as reinforcers (Roscoe et al., 1999); as a result, we felt that nonapproach to a stimulus presented singly would yield more valid information about lack of preference. Each participant was allowed to sample each item prior to the start of the assessment. Each stimulus was placed singly in front of the participant for 5 s; contingent on an approach response, the participant was allowed to consume the item. If no approach response occurred within 5 s, the item was removed and a new trial began. Five trials were conducted with each stimulus, and three stimuli that were never consumed were used as the NP items.

Experimental Conditions

Target behaviors during the reinforcer assessment were selected for each participant using criteria similar to those used in Study 1. If necessary, brief training on the target response occurred until the participant emitted the response independently.

Baseline. The purpose of this condition was to determine whether the participant would engage in the response in the absence of any contingencies. No consequences were delivered for engaging in the target response.

Nonpreferred stimuli (NP). The purpose of this condition was to determine whether the NP items would function as reinforcers. Contingent on the target response, one NP item was delivered on an FR 1 schedule. Each item was assessed separately in a session, and sessions continued until a stable trend was established. Participants were prompted to engage in the target response prior to the start of the session to establish contact with the contingency.

Varied nonpreferred stimuli (VNP). If no or little responding was observed in the previous condition, the three NP stimuli were presented in a varied manner throughout the session. One NP edible item was delivered contingent on the target response on an FR 1 schedule; the order of delivering the three NP items was constant within a session but was randomized across sessions. The participant was prompted to engage in the target response three times prior to the session.

VNP + HQ. This condition was identical to the previous one, except that the HQ stimulus was added to the arrangement. One of the four edible items (one HQ and three NP) was delivered contingent on the target response (again, order was constant within a session but was randomized across sessions). The participant was prompted to engage in the target response four times prior to session.

HQ alone. Participants were exposed to one of two types of conditions in which the HQ stimulus was delivered alone, depending on their performance in previous conditions. First, if no increase in responding had been observed during any previous condition (NP, VNP, VNP + HQ), the HQ stimulus was delivered on an FR 1 schedule simply to determine whether a reinforcement effect, which was absent to that point, could be obtained. Second, if responding in the VNP + HQ condition was higher than in the VNP condition, it was possible that improved performance was solely a function of the HQ stimulus. To control for this possibility, the HQ stimulus was delivered on an FR 4 schedule, which matched the schedule of HQ delivery in the VNP + HQ condition. The participant was prompted to engage in the target response either once (FR 1) or four times (FR 4) prior to session.

Experimental Design

The conditions were presented in the following order: baseline, NP, VNP, VNP + HQ, and HQ alone, although some modifications were made in the sequence based on performance. A reversal design was used to demonstrate experimental control.

Results and Discussion

Figure 4 shows results from the two preference assessments. The item approached

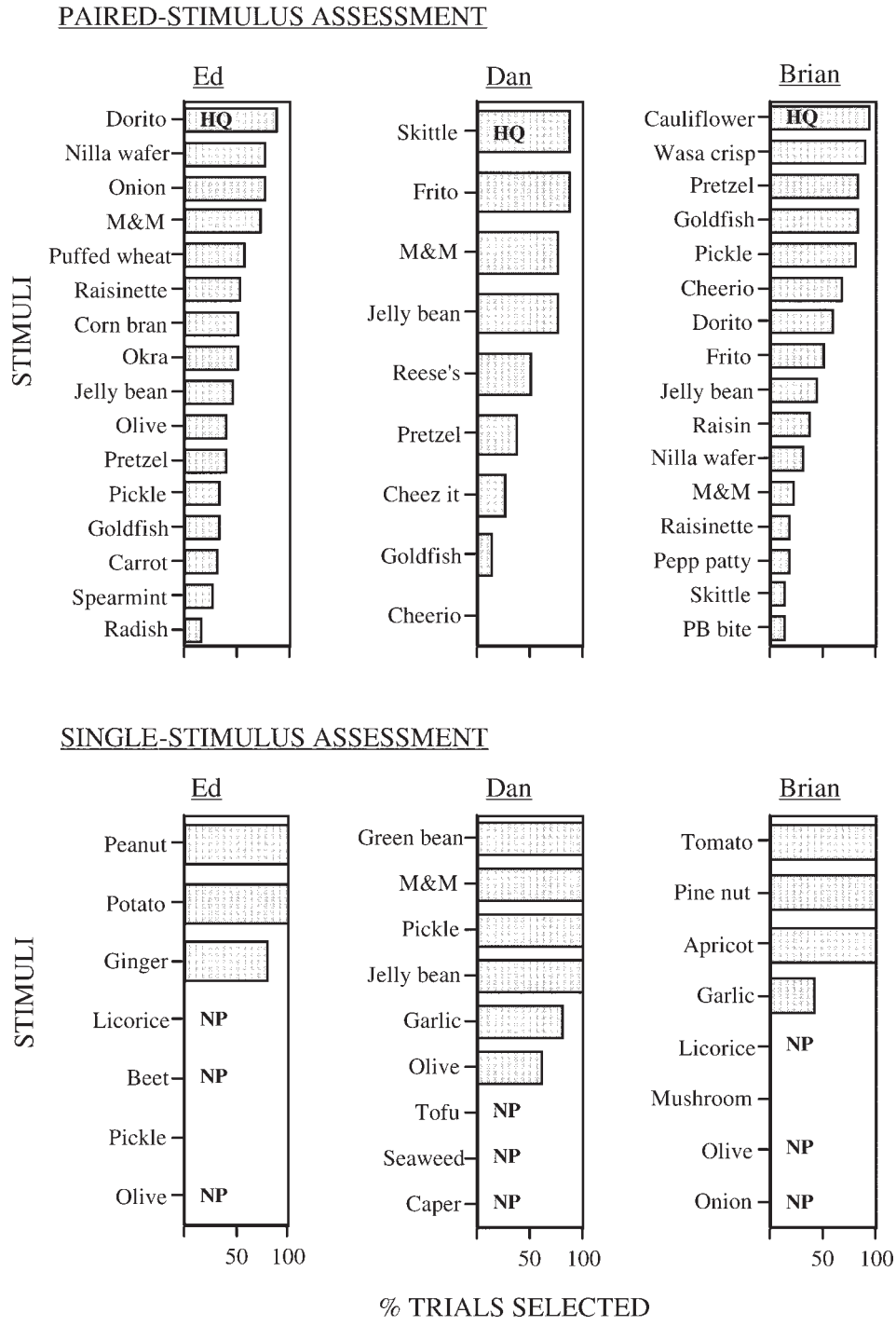

Figure 4. Results of paired-stimulus and single-stimulus preference assessments. Items designated as HQ and nonpreferred (NP) were used in Study 2.

most frequently in the paired-stimulus assessment was designated as the HQ stimulus; three items that were never approached in the single-stimulus assessment were designated as NP stimuli.

Figure 5 shows results from the reinforcer assessments. Ed never engaged in the target response during baseline or during any of the subsequent conditions in which the NP stimuli were present (NP, VNP, and VNP + HQ). Thus, it appeared that none of the stimuli, including the HQ item, were reinforcers. We then implemented a condition in which the HQ stimulus was delivered on an FR 1 schedule and observed an immediate increase in responding. A return to the VNP + HQ condition resulted in a return to zero responding. These results suggest that the NP items not only failed to serve as reinforcers but also obscured the reinforcing effects of the HQ item. In Ed's case, variation of NP and HQ items had an overall detrimental effect on performance.

Dan exhibited no responses during baseline, NP, and VNP conditions. His responding increased immediately during the subsequent VNP + HQ condition. We therefore implemented the HQ (FR 4) condition and observed no change in response rates. When the VNP + HQ condition was reinstated, Dan's performance was variable initially. He exhibited no responses during two sessions but subsequently responded at rates similar to the two previous conditions. His responding was maintained during the final HQ (FR 4) condition. Dan's results differed from Ed's in that the inclusion of a highly preferred item with nonpreferred items in a varied arrangement (VNP + HQ) appeared to facilitate performance. However, when these elements were presented separately, Dan did not respond during the VNP condition, whereas his response rates during the HQ (FR 4) condition were similar to those observed during the VNP + HQ condition. In addition, he consumed only one of the NP items (capers) throughout the VNP + HQ condition. Thus, it seems likely that Dan's responding throughout the study can be attributed to the presence of the HQ item.

Brian responded sporadically during the baseline and NP conditions. His responding also was low during the VNP condition; however, unlike his response pattern in previous conditions, he exhibited some responses during every session. During the VNP + HQ condition, Brian's responding increased noticeably but remained somewhat variable, and continued in a similar manner (although at a slightly lower rate) during a return to the VNP condition. Given the modest improvement observed in his performance during the VNP + HQ condition and the fact that his performance also showed a slight improvement during the VNP condition relative to the NP condition, we chose to reexamine the NP and VNP conditions rather than to implement the HQ (FR 4) condition. Brian's responding decreased to zero during the return to baseline, increased slightly during the NP condition, and increased further during the final VNP condition. His results may illustrate two effects related to stimulus variation. First, the inclusion of an HQ stimulus may have enhanced the reinforcing effects of nonpreferred stimuli in the VNP + HQ condition, although the extent to which observed increases in responding were solely a function of the HQ stimulus was unknown because we did not implement the HQ (FR 4) condition. Second, the VNP condition resulted in a modest improvement in performance over baseline and NP conditions.

GENERAL DISCUSSION

We examined several characteristics of stimulus variation in the context of reinforcement facilitation under conditions of concurrent and single reinforcement. In Study 1, 8 participants were exposed to conditions in which the reinforcing effects of a stimulus ranked at the top of a preference hierarchy (HQ) were compared with those of varied stimuli ranked

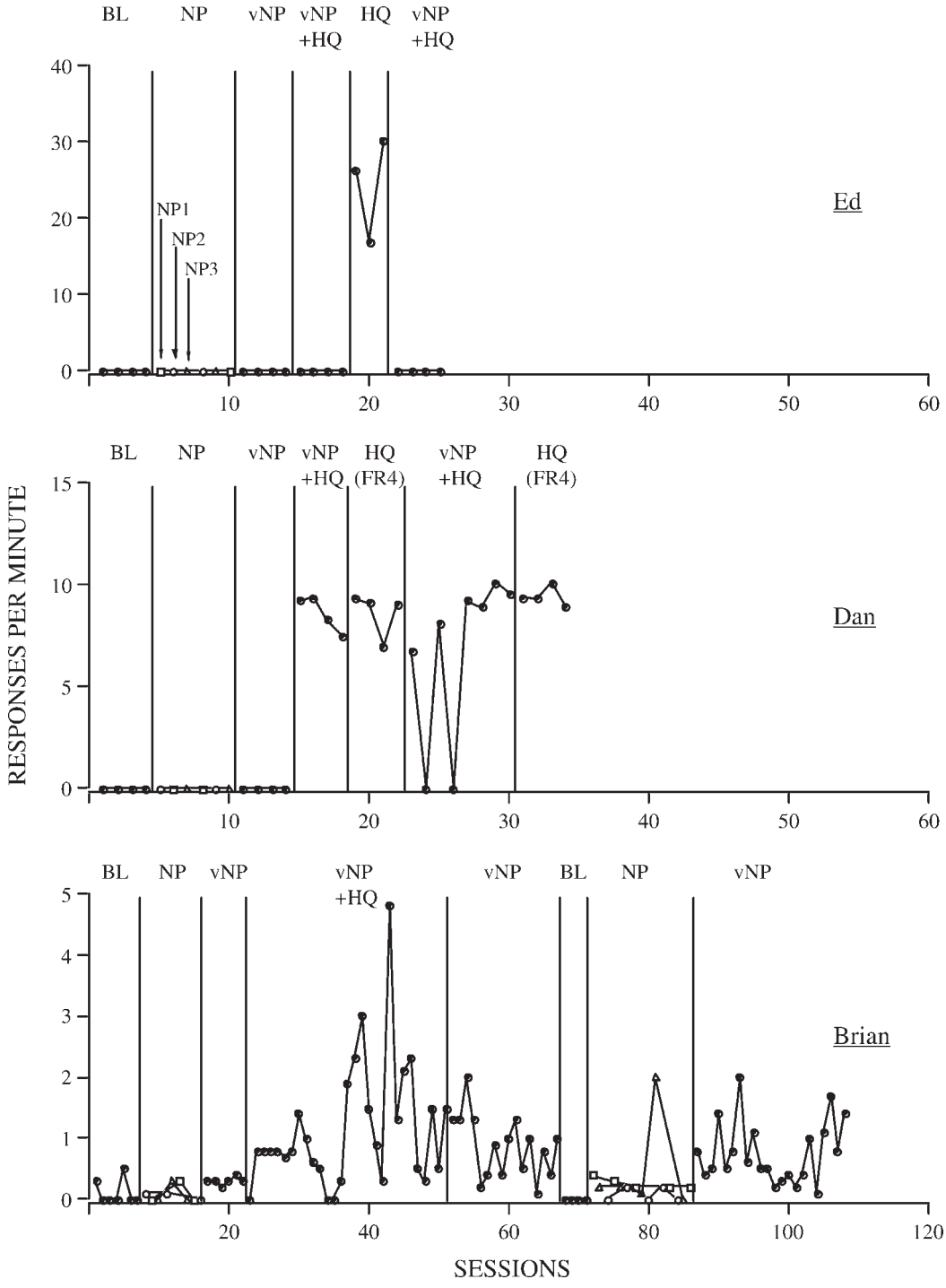

Figure 5. Results of reinforcer assessments for Ed, Dan, and Brian.

at the bottom (VLQ) under a concurrent-reinforcement arrangement. Seven of the 8 participants showed exclusive preference for the single HQ stimulus; 1 participant (Josh) switched between reinforcement options. When the HQ stimulus was compared with varied stimuli ranked second, third, and fourth in the preference hierarchy (VSLQ), half the participants showed preference for the VSLQ stimuli; the other half switched preference across sessions. These data suggest that variation does little to enhance the effects of NP stimuli, although this conclusion requires several qualifications.

First, it is not clear that competition between the VSLQ stimuli and the HQ stimulus can be attributed to variation per se because, as previously noted, the single presentation of any one (or maybe all) of the SLQ stimuli could have produced similar results. We did not conduct HQ versus single SLQ comparisons because our main interest was in the effects of LQ stimuli. When we observed that the VLQ condition produced negligible results, we included the VSLQ condition to more closely approximate conditions of the Bowman et al. (1997) study and observed similar effects. Thus, results of the concurrent-schedule comparisons showed that stimulus variation is unlikely to improve competition between stimuli ranked high and low on a preference hierarchy; the effects of variation using more similarly ranked stimuli is less clear. Results of the HQ versus VSLQ comparison do show, however, that stimulus variation may allow SLQ stimuli to compete effectively with one HQ stimulus when both options are simultaneously available.

Second, concurrent presentation of stimuli whose preference is established via paired-stimulus assessments is likely to produce false-negative outcomes. For example, Roscoe et al. (1999) showed that, although stimuli never selected in a paired-stimulus assessment did not compete with highly preferred stimuli under a concurrent-schedule arrangement, the same stimuli functioned as reinforcers under a single-schedule arrangement. Therefore, negative conclusions about the facilitative effects of stimulus variation with NP stimuli in Study 1 must be limited to situations that involve reinforcer competition. In most applied settings, however, therapists would want to know whether stimulus variation facilitates *any* reinforcement effect. Results of Egel's (1980, 1981) studies indicated that variation enhances the reinforcing effects of stimuli that already function as reinforcers; the unknown implication of these results, as well as those from Study 1, is whether variation might produce a reinforcement effect with stimuli that otherwise do not function as reinforcers.

We therefore used a different methodology to evaluate stimulus variation in Study 2. NP stimuli were selected based on the absence of approach during single-stimulus preference assessments, because results of previous research (Green et al., 1988; Pace et al., 1985) showed that such stimuli do not function as reinforcers; this was verified during the single NP conditions, in which 2 participants (Ed and Dan) never responded, and 1 participant (Brian) responded at rates similar to those observed during baseline. When the VNP condition was implemented, Ed and Dan continued to exhibit no responses, whereas Brian's responding increased, although only slightly. Thus, as might be expected, varying the delivery of nonpreferred stimuli appeared to have negligible effects on behavior. We then evaluated the effects of including the HQ stimulus as one of the items in the varied arrangement (VNP + HQ). Ed continued to exhibit no responses, and his performance during the subsequent HQ condition suggested that the inclusion of VNP items in the VNP + HQ condition made his behavior less sensitive to the effects of the HQ stimulus. In other words, varying NP and HQ items produced detrimental results. One interpretation of Ed's results is that the VNP + HQ condition was functionally the equivalent of

delivering the HQ stimulus on an FR 4 schedule that did not maintain responding (i.e., extinction occurred). Another interpretation is that delivery of the NP items served as punishment for responding. By contrast, Dan's responding increased immediately under the VNP + HQ condition, suggesting a facilitative effect. However, it appeared that this effect could be attributed solely to the presence of the HQ stimulus because Dan rarely consumed the NP items, and his responding was maintained when the VNP items were eliminated in the HQ (FR 4) condition. Finally, Brian's responding during the VNP + HQ condition, although higher than during previous conditions, seemed modest at best and prompted us to reexamine the VNP arrangement. When the VNP condition was reimplemented, Brian's rate of responding was higher than during two previous baselines and two NP conditions and was similar to that observed during the VNP + HQ condition. Thus, although it was an unusual finding, the varied delivery of NP stimuli seemed to increase the reinforcing effects of those stimuli on Brian's behavior.

The applied relevance of Brian's data may be limited because it is unlikely that therapists would deliver as reinforcers those stimuli to which approach had never been observed. However, Ivancic (2000) has noted that situations may arise in which individuals approach few, if any, stimuli in a preference assessment. In such situations, varying the delivery of any available stimuli, or including a single preferred stimulus in a varied arrangement with those stimuli, expands the range of available reinforcement options.

In summary, results of the present studies indicate that stimulus variation has little effect on the reinforcing capability of NP stimuli relative to that of a highly preferred stimulus when both alternatives are available. Different results, however, might be obtained with stimulus variation in the context of a single-task arrangement. Brian's data suggested that the varied delivery of NP stimuli or the inclusion of a preferred stimulus in a varied arrangement with NP stimuli may produce modest improvements in performance over that observed when NP stimuli are delivered singly. Furthermore, several arrangements not examined in Study 2 might yield even clearer results. For example, Brian's VNP condition was conducted for relatively few sessions; it is possible that longer exposure to this condition would be associated with further increases in responding. In addition, a more optimal arrangement for increasing the effectiveness of NP stimuli might consist of pairing them singly with a highly preferred stimulus in an attempt to establish the former as a reinforcer (e.g., see Hanley, Iwata, Roscoe, Thompson, & Lindberg, 2003). Finally, a less optimistic conclusion about the effects of stimulus variation was revealed in Ed's data, which indicated that varying nonpreferred and preferred stimuli as consequences may occasionally produce undesirable outcomes.

REFERENCES

Bowman, L. G., Piazza, C. C., Fisher, W. W., Hagopian, L. P., & Kogan, J. S. (1997). Assessment of preference for varied versus constant reinforcers. *Journal of Applied Behavior Analysis*, *30*, 451–458.

DeLeon, I. G., & Iwata, B. A. (1996). Evaluation of a multiple-stimulus presentation format for assessing reinforcer preferences. *Journal of Applied Behavior Analysis*, *29*, 519–533.

Egel, A. L. (1980). The effects of constant vs. varied reinforcer presentation on responding by autistic children. *Journal of Experimental Child Psychology*, *30*, 455–463.

Egel, A. L. (1981). Reinforcer variation: Implications for motivating developmentally disabled children. *Journal of Applied Behavior Analysis*, *14*, 345–350.

Fisher, W. W., Piazza, C. C., Bowman, L. G., Hagopian, L. P., Owens, J. C., & Slevin, I. (1992). A comparison of two approaches for identifying reinforcers for persons with severe to profound disabilities. *Journal of Applied Behavior Analysis*, *25*, 491–498.

Green, C. W., Reid, D. H., White, L. K., Halford, R. C., Brittain, D. P., & Gardner, S. M. (1988). Identifying reinforcers for persons with profound handicaps: Staff opinion versus systematic assessment of preferences. *Journal of Applied Behavior Analysis*, *21*, 31–43.

Hanley, G. P., Iwata, B. A., Roscoe, E. M., Thompson, R. H., & Lindberg, J. S. (2003). Response-restriction analysis: II. Alteration of activity preferences. *Journal of Applied Behavior Analysis*, *36*, 59–76.

Herrnstein, R. J., & Loveland, D. H. (1975). Maximizing and matching on concurrent ratio schedules. *Journal of the Experimental Analysis of Behavior*, *24*, 107–116.

Ivancic, M. T. (2000). Stimulus preference and reinforcer assessment applications. In J. Austin, & J. E. Carr (Eds.), *Handbook of applied behavior analysis* (pp. 19–38). Reno, NV: Context Press.

North, S. T., & Iwata, B. A. (2005). Motivational influences on performance maintained by food reinforcement. *Journal of Applied Behavior Analysis*, *38*, 317–333.

Pace, G. M., Ivancic, M. T., Edwards, G. L., Iwata, B. A., & Page, T. J. (1985). Assessment of stimulus preference and reinforcer value with profoundly retarded individuals. *Journal of Applied Behavior Analysis*, *18*, 249–255.

Roscoe, E. M., Iwata, B. A., & Kahng, S. (1999). Relative versus absolute reinforcement effects: Implications for preference assessments. *Journal of Applied Behavior Analysis*, *32*, 479–493.

Received July 21, 2004
Final acceptance June 30, 2005
Action Editor, Iser DeLeon